



Nounu, A., Walker, V., & Richmond, R. C. (2021). Letter regarding, "Association between the use of aspirin and risk of lung cancer results from pooled cohorts and Mendelian randomization analyses". *Journal of Cancer Research and Clinical Oncology*.
<https://doi.org/10.1007/s00432-020-03508-z>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1007/s00432-020-03508-z](https://doi.org/10.1007/s00432-020-03508-z)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Springer Nature at <https://doi.org/10.1007/s00432-020-03508-z>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Letter regarding, “Association between the use of aspirin and risk of lung cancer: results from pooled cohorts and Mendelian randomization analyses”

Aayah Nounu¹ · Venexia Walker¹ · Rebecca C. Richmond¹

Received: 23 November 2020 / Accepted: 17 December 2020
© The Author(s) 2021

Dear Editor,

We are writing to comment on the paper “Association between the use of aspirin and risk of lung cancer: results from pooled cohorts and Mendelian randomization analyses” by Jiang et al. (Jiang et al. 2020). Findings from this Mendelian randomization (MR) study suggest that aspirin use decreases the incidence of lung cancer, specifically overall lung cancer and squamous cell carcinoma.

Observational studies have provided some evidence for the use of aspirin as a chemopreventive agent in lung cancer (relative risk 0.93, 95%CI 0.87–1.00) (Qiao et al 2018). However, few randomized trials have been carried out to answer this question. Using genetic variation as a method of randomization to aspirin use and testing for association with specific cancers within an MR framework is, therefore, an attractive method to appraise causality before long and costly trials are conducted.

The authors used genetic variants from a genome-wide association study (GWAS) on aspirin use conducted by the Neale Lab (UK Biobank—Neale lab) to test whether aspirin intake was causally related to lung cancer incidence. However, there are some potential concerns with this method of instrumenting aspirin use.

One of the first major concerns with using instruments from a GWAS of drug use is disentangling the genetic variants for the drug from those for the drug’s indication. Using the MRC IEU OpenGWAS database (Elsworth et al. 2020), we found that the SNPs that predict aspirin use by this study (rs583104, rs2521501, rs10455872, rs73015016, rs7412, rs1831733, rs117733303) have all been shown to be

associated with coronary artery disease (CAD) risk at least at genome-wide significance (van der Harst and Verweij 2018; Nikpay et al. 2015). Therefore, it may be that these SNPs increase the risk of CAD and that these individuals are being instructed to take aspirin as a preventative measure, thereby confounding the SNP association with aspirin use.

Many of the SNPs used to instrument aspirin use are also associated with a large number of other risk factors in the MRC IEU OpenGWAS database (Elsworth et al. 2020). This raises the potential for the violation of two of the MR assumptions: no confounding (independence assumption) and no horizontal pleiotropy (exclusion restriction assumption). Specifically, SNPs rs583104, rs10455872, rs73015016, rs7412 and rs117733303 have previously been associated with levels of low-density lipoprotein (LDL) cholesterol at genome-wide significance (Global Lipids Genetics Consortium, 2013; Prins et al. 2017), with increasing LDL-cholesterol levels leading to increased risk of coronary heart disease (Richardson et al. 2020). If cardiovascular risk factors are causally related to lung cancer incidence, as suggested by a previous MR study where increasing LDL-cholesterol levels was inversely associated with lung cancer incidence (OR 0.90, 95% CI 0.84–0.97 per SD of 38 mg/dl) (Carreras-Torres et al. 2017), then this may introduce confounding into the MR analysis.

Alternatively, associations between the genetic variants being used to instrument aspirin use and other risk factors could indicate violation of the exclusion restriction assumption of MR—namely, that the SNP is only affecting the outcome via the exposure of interest (Lawlor et al. 2008). When SNPs are also associated with other risk factors that may affect disease risk, this is termed horizontal pleiotropy (Burgess and Thompson 2013); however, we do acknowledge that the MR Egger regression was conducted and found little evidence of pleiotropy (Bowden et al. 2015). Furthermore, a weighted median approach was carried out and consistent results with the IVW were observed for overall lung cancer

✉ Aayah Nounu
an0435@bristol.ac.uk

¹ Medical Research Council (MRC) Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK

Table 1 MR results using exposure instruments transformed to the log odds scale

	IVW				Weighted median				MR Egger			
	OR	LCI	UCI	<i>P</i> value	OR	LCI	UCI	<i>P</i> value	OR	LCI	UCI	<i>P</i> value
Lung cancer	0.69	0.51	0.94	0.02	0.60	0.42	0.87	0.01	0.36	0.16	0.77	0.05
Lung adenocarcinoma	0.71	0.38	1.33	0.29	0.55	0.31	0.98	0.04	0.38	0.06	2.22	0.33
Squamous cell lung cancer	0.48	0.27	0.87	0.02	0.50	0.27	0.93	0.03	0.36	0.06	2.03	0.30

(OR 1.32×10^{-4} , 95% CI 1.69×10^{-7} to 0.10, *P* value: 0.05) indicating a similar causal effect even if 50% of the weight came from invalid instruments (Bowden et al. 2015).

Another threat to the validity of findings is the potential for selection bias. Since many of the SNPs may also be proxying liability to CAD, this may indicate survival bias, whereby individuals with higher risk of CAD die prematurely and, therefore, do not live long enough to be diagnosed with lung cancer. A frailty analysis could be conducted in this case to re-estimate the causal estimate in the presence of survival bias (Noyce et al. 2017).

A final concern with the results presented are the very large effect sizes (overall lung cancer, OR 0.042, 95% CI 0.003–0.564 and squamous cell lung cancer, OR 0.002, 95% CI 1.21×10^{-5} to 0.301) obtained from the MR analysis, which are much larger in magnitude than the corresponding observational estimates presented in the study (overall lung cancer relative risk (RR): 0.95, 95% CI 0.91–0.98, *P* value: 0.004; and squamous cell lung cancer RR: 0.80, 95% CI: 0.65 to 0.98, *P* value: 0.034). While it is possible for MR to estimate larger causal effects than corresponding observational analyses; for example, in the presence of negative confounding or short-term exposure, a more detailed assessment of the genetic estimates used in the analysis is first required. In particular, the GWAS for aspirin use conducted by the Neale lab was conducted on the absolute risk difference rather than log-odds scale and, therefore, the MR effect estimates presented are unlikely to be directly comparable to those obtained in the accompanying observational analysis.

Using the same SNPs and datasets, we first replicated the results to confirm that the same exposure and outcome datasets were being used before transforming the SNP–exposure associations to a more interpretable scale for binary traits (the log odds scale) using the formula provided by Elsworth et al. (2019) and re-conducting the MR analysis (code to reproduce the analysis in this paper: [https://github.com/an0435/aspirin_lung_cancer_r_MR/](https://github.com/an0435/aspirin_lung_cancer_MR/)). The transformed MR results can be interpreted as the OR for lung cancer per doubling of aspirin use. After conversion, a doubling of aspirin use decreases the risk of lung cancer and squamous cell lung cancer by 31% and 52%, respectively (lung cancer IVW OR 0.69, 95% CI 0.51–0.94 and squamous cell lung cancer IVW OR 0.48, 95% CI 0.27–0.87) (Table 1). Using the numbers

of cases and sample sizes from the cohort studies listed in the paper, the prevalence of lung cancer was 0.51% (72,782/14,369,951*100). When the disease is rare in a population such as with lung cancer (prevalence below 10%), the OR can be interpreted as a relative risk, making the results from the observational analysis and MR comparable (Sedgwick, 2014). These transformed MR results and their confidence intervals now partially overlap with the observational estimates. (overall lung cancer observational RR: 0.95, 95% CI 0.91–0.98, MR OR 0.69, 95% CI: 0.51–0.94; squamous cell lung cancer observational RR: 0.80, 95% CI 0.65–0.98, MR OR 0.48, 95% CI 0.27–0.87).

One alternative approach to study drug effects using MR is to identify SNPs that mimic a drug's mechanism of action by investigating SNPs in the genes of the targeted protein (Gill et al. 2019). For example, statins inhibit the enzyme 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGCR) resulting in reduced levels of LDL-cholesterol (FERENCE et al. 2015). Based on this understanding, SNPs that are in or around (within 100 kb) the *HMGCR* gene and that are associated with LDL-cholesterol have proven a useful method to instrument exposure to statins in MR studies (FERENCE et al. 2015, 2016). In the case of drugs such as aspirin that have multiple targets, it may be useful to conduct proteomic analysis to identify proteins targeted by the drug and instrument the effect of changes in mRNA/protein levels on cancer risk (Nounu et al. 2020). Instrumenting levels of mRNA/protein expression provides a continuous exposure, compared to aspirin use which is a binary variable and, therefore, results in complications when conducting and interpreting MR analyses (Burgess and Labrecque 2018).

Whilst we acknowledge that conducting MR studies of drug use and cancer incidence would provide much needed answers for clinical intervention, careful consideration in MR study design is needed when instrumenting drug use in MR to avoid the potential pitfalls highlighted. Methods and guidelines are now available for informing best practice in MR (Burgess et al. 2020; Davey Smith et al. 2019) so that appropriate inference can be made.

Acknowledgements The authors thank Dr Philip C. Haycock for his feedback considering the manuscript.

Author contributions AN provided the initial concept, performed the statistical analysis and drafted the manuscript. VW helped with reviewing and editing the manuscript. RCR helped with supervising as well as reviewing and editing the manuscript.

Funding This work was funded by Cancer Research UK (C18281/A19169) and the Medical Research Council Integrative Epidemiology Unit, which is supported by the Medical Research Council and the University of Bristol [MC_UU_00011/1 and MC_UU_00011/4].

Code availability The code to reproduce this analysis can be found at (https://github.com/an0435/aspirin_lung_cancer_MR/).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Availability of data and material Data are freely available on the online platform for MR Base (<http://app.mrbase.org/>) as well as through the TwoSampleMR R package (github.com/MRCIEU/TwoSampleMR).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bowden J, Davey Smith G, Burgess S (2015) Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol* 44:512–525
- Burgess S, Labrecque JA (2018) Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol* 33:947–952
- Burgess S, Smith GD, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, Theodoratou E (2020) Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res* 4:186
- Burgess S, Thompson SG (2013) Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol* 42:1134–1144
- Carreras-Torres R, Johansson M, Haycock PC, Wade KH, Relton CL, Martin RM, Davey Smith G, Albanes D, Aldrich MC, Andrew A, Arnold SM, Bickeböller H, Bojesen SE, Brunström H, Manjer J, Brüske I, Caporaso NE, Chen C, Christiani DC, Christian WJ et al (2017) Obesity, metabolic factors and risk of different histological types of lung cancer: a Mendelian randomization study. *PLoS ONE* 12:e0177875
- Davey Smith G, Davies NM, Dimou N, Egger M, Gallo V, Golub R, Higgins JP, Langenberg C, Loder EW, Richards JB, Richmond R, Skrivankova VW, Swanson S, Timpson NJ, Tybjaerg-Hansen A, VanderWeele TJ, Woolf BA, Yarmolinsky J (2019) STROBE-MR: Guidelines for strengthening the reporting of Mendelian randomization studies. *PeerJ*. <https://peerj.com/preprints/27857/>
- Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, Bates P, Palmer T, Haberland V, Davey Smith G, Zheng J, Haycock P, Gaunt TR, Hemani G (2020) The MRC IEU OpenGWAS data infrastructure. *bioRxiv*. <https://doi.org/10.1101/2020.08.10.244293v1>
- Elsworth B, Mitchell R, Raistrick C, Paternoster L, Hemani G & Gaunt T (2019) MRC IEU UK Biobank GWAS pipeline, version 2, 18/01/2019. <https://data.bris.ac.uk/datasets/pnoat8cxo0u52p6ynfaekeigi/MRCIEUUKBiobankGWASpipelineversion2.pdf>. Accessed 16 Nov 2020
- Ference BA, Majeed F, Penumetcha R, Flack JM, Brook RD (2015) Effect of naturally random allocation to lower low-density lipoprotein cholesterol on the risk of coronary heart disease mediated by polymorphisms in NPC1L1, HMGCR, or Both: A 2 × 2 factorial mendelian randomization study. *J Am Coll Cardiol* 65:1552–1561
- Ference BA, Robinson JG, Brook RD, Catapano AL, Chapman MJ, Neff DR, Voros S, Giugliano RP, Davey Smith G, Fazio S, Sabatine MS (2016) Variation in PCSK9 and HMGCR and risk of cardiovascular disease and diabetes. *N Engl J Med* 375:2144–2153
- Gill D, Georgakis MK, Koskeridis F, Jiang L, Feng Q, Wei W-Q, Theodoratou E, Elliott P, Denny JC, Malik R, Evangelou E, Dehghan A, Dichgans M, Tzoulaki I (2019) Use of genetic variants related to antihypertensive drugs to inform on efficacy and side effects. *Circulation* 140:270–279
- Global Lipids Genetics Consortium (2013) Discovery and refinement of loci associated with lipid levels. *Nat Genet* 45:1274–1283
- van der Harst P, Verweij N (2018) Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res* 122:433–443
- Jiang Y, Su Z, Li C, Wang R, Wen Y, Liang H, He J, Liang W (2020) Association between use of aspirin and risk of lung cancer: results from pooled cohorts and mendelian randomization analyses. *J Cancer Res Clin Oncol*. <https://doi.org/10.1007/s00432-020-03394-5>
- Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G (2008) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med* 27:1133–1163
- Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I et al (2015) A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 47:1121–1130
- Nounu A, Greenhough A, Heesom KJ, Richmond RC, Zheng J, Weinstein SJ, Albanes D, Baron JA, Hopper JL, Figueiredo JC, Newcomb PA, Lindor NM, Casey G, Platz EA, Marchand L Le, Ulrich CM, Li CI, Duijnhoven FJ van, Gsur A, Campbell PT, et al (2020) A combined proteomics and Mendelian randomization approach to investigate the effects of aspirin-targeted proteins on colorectal cancer. <https://cebp.aacrjournals.org/content/early/2020/12/12/1055-9965.EPI-20-1176>
- Noyce AJ, Kia DA, Hemani G, Nicolas A, Price TR, Pablo-Fernandez E De, Haycock PC, Lewis PA, Foltynie T, Smith GD, International Parkinson Disease Genomics Consortium, Schrag A, Lees AJ, Hardy J, Singleton A, Nalls MA, Pearce N, Lawlor DA, Wood NW (2017) Estimating the causal influence of body mass index on risk of Parkinson disease: a Mendelian randomisation study. *PLoS Med* 14:e1002314
- Prins BP, Kuchenbaecker KB, Bao Y, Smart M, Zabaneh D, Fatemifar G, Luan J, Wareham NJ, Scott RA, Perry JRB, Langenberg C, Benzeval M, Kumari M, Zeggini E (2017) Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. *Sci Rep* 7:11008
- Qiao Y, Yang T, Gan Y, Li W, Wang C, Gong Y, Lu Z (2018) Associations between aspirin use and the risk of cancers: a meta-analysis of observational studies. *BioMed Cent Cancer* 18:1–57
- Richardson TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, Holmes MV (2020) Evaluating the relationship

between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: a multivariable Mendelian randomisation analysis. *PLOS* 17:e1003062

Sedgwick P (2014) Relative risks versus odds ratios. *BMJ* 348:g1407
UK Biobank—Neale lab. <http://www.nealelab.is/uk-biobank/>. Accessed 19 Oct 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.